Review

Structure Prediction: New Insights into Decrypting Long Noncoding RNAs

Kun Yan [1], Yasir Arfat [1], Dijie Li [1], Fan Zhao [1], Zhihao Chen [1], Chong Yin [1], Yulong Sun [1], Lifang Hu [1], Tuanmin Yang [2],[†] and Airong Qian [1],*

[1] Key Laboratory for Space Bioscience & Biotechnology, Institute of Special Environmental Biophysics, School of Life Sciences, Northwestern Polytechnical University, 127 Youyi Xilu, Xi'an 710072, China; yan@mail.nwpu.cn (K.Y.); Yasir@mail.nwpu.edu.cn (Y.A.); lidijie@mail.nwpu.edu.cn (D.L.); sofan@mail.nwpu.edu.cn (F.Z.); chzhh@mail.nwpu.edu.cn (Z.C.); Yinchong42@gmail.com (C.Y.); yulongsun@nwpu.edu.cn (Y.S.); lifanglinn@mail.nwpu.edu.cn (L.H.)
[2] Department of Bone Disease Oncology, Hong-Hui Hospital, Xi'an Jiaotong University College of Medicine, South Door slightly Friendship Road 555, Xi'an 710054, China; doctoryangtm@163.com
* Correspondence: qianair@nwpu.edu.cn; Tel./Fax: +86-29-8849-1840
† Co-corresponding author.

Abstract: Long noncoding RNAs (lncRNAs), which form a diverse class of RNAs, remain the least understood type of noncoding RNAs in terms of their nature and identification. Emerging evidence has revealed that a small number of newly discovered lncRNAs perform important and complex biological functions such as dosage compensation, chromatin regulation, genomic imprinting, and nuclear organization. However, understanding the wide range of functions of lncRNAs related to various processes of cellular networks remains a great experimental challenge. Structural versatility is critical for RNAs to perform various functions and provides new insights into probing the functions of lncRNAs. In recent years, the computational method of RNA structure prediction has been developed to analyze the structure of lncRNAs. This novel methodology has provided basic but indispensable information for the rapid, large-scale and in-depth research of lncRNAs. This review focuses on mainstream RNA structure prediction methods at the secondary and tertiary levels to offer an additional approach to investigating the functions of lncRNAs.

Keywords: lncRNAs; function; structure prediction; secondary structure; tertiary structure

1. Introduction

The term noncoding RNAs (ncRNAs) refers to RNA transcripts that do not encode proteins [1]. Approximately 93% of human genomic DNA can be transcribed into RNAs [2]. Merely 2% of these RNAs will be translated into approximately 20,000 types of protein translation products, and the remaining 98% of these RNAs represent noncoding RNAs that are rarely translated [3]. ncRNAs can be classified as housekeeping or regulatory ncRNAs. Housekeeping ncRNAs include ribosome RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNA (snRNAs) and small nucleolar RNA (snoRNAs), while regulatory RNAs include small interfering RNA (siRNAs), microRNA (miRNAs), long noncoding RNAs (lncRNA), piRNA, natural antisense transcripts (NATs) and circular RNA (circRNAs) [4–7].

Except for tRNAs and rRNAs, ncRNAs have been traditionally disregarded as "transcriptional noise" [8]. Although proteins have long been considered to carry genetic information, emerging evidence implies that ncRNAs are also involved in the regulation of gene expression that impacts the growth and development of organisms [9–11]. Compared with short RNAs (<200 nt), highly

transcribed long noncoding RNAs (lncRNAs) (>200 nt) may perform more complex biological functions [12–14]. These RNAs have been implicated in the regulation of gene expression at the transcriptional or posttranscriptional level exerting effects on dosage compensation, chromatin regulation, genomic imprinting, nuclear organization, alternative splicing of pre-mRNA and many other biological processes [15–17]. Considering the participation of lncRNAs in various aspects of gene expression affecting the differentiation and development of organisms, it is not surprising that the dysregulation of lncRNAs has been involved in disease [18,19]. According to a genome-wide association study, 43% of reported trait/disease-associated SNPs (TASs) were intergenic, suggesting essential roles for ncRNAs in common diseases [20]. Furthermore, Chen *et al.* [21] created lncRNADisease, a database of 166 lncRNA-associated diseases. lncRNADisease collected nearly 480 entries of experimentally validated lncRNA-disease associations. The recognition of the important roles of lncRNAs in human disease has provided novel diagnostic and therapeutic opportunities [22].

Given the wide range of biological functions in which lncRNAs have been implicated, we predict that many more lncRNAs will be determined to have important functions. For many RNAs, there is a close relationship between structure and function [23–25]. Their structural diversity allows for RNA to perform various functions, including catalytic, organizational and other regulatory functions [26,27]. Generating structural models of these RNAs that are faithful to their native structures is essential because the structure of RNA influences its transcription, splicing, cellular localization, translation and turnover [28]. Thus, acquiring structural information for RNA is often the first step towards exploring its function [29].

2. Review

This review focuses on lncRNAs, which comprise the least understood class of ncRNAs. Their functions, mechanisms, roles in epigenetics and relationships with diseases are introduced. Moreover, ncRNA structure prediction methods such as Foldalign [30], Pfold [31], Mfold [32], RNAfold [33], RNAshapes [34], RNAstructure [35], NAST [36], iFoldRNA [37], and 3dRNA [38] are reviewed (Figure 1). Furthermore, the theories underlying each method as well as the advantages and pitfalls of their applications are provided. Based on this summary, another step in the understanding of lncRNAs can be achieved. As the secondary/tertiary structures of several functionally understood lncRNAs have been predicted (or experimentally verified), RNA structure predictions may help identify additional functional lncRNAs and may thus offer clues for the design of targeted small molecule therapeutics to promote drug development and the treatment of diseases [39].
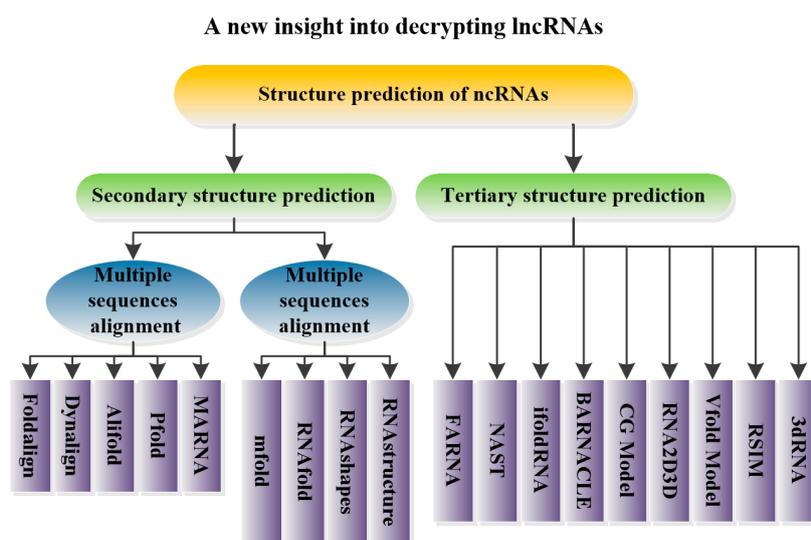
Figure 1. The graphical abstract of this review.

3. ncRNAs

Recent advances in the study of ncRNAs have demonstrated the existence of ncRNAs in a wide variety of mammalian transcriptomes. miRNAs, siRNAs, lncRNAs and other ncRNAs have been found to play significant roles in various physiological and developmental processes in eukaryotes, furthering our knowledge of ncRNAs.

3.1. Evolutionary Conservation of ncRNAs

A plethora of analyses have revealed that the majority of ncRNAs possess low rates of interspecies conservation at the primary sequence level as protein-coding RNAs [40]. Analyses of the conservation of various classes of ncRNAs have revealed that these RNAs differ in their evolutionary conservation. Of these ncRNA transcripts, miRNAs and snoRNAs are highly sequence-conserved [41,42], even more so than mRNAs or other protein-coding transcripts [43]. Conversely, longer ncRNAs have been found to evolve at a rapid rate, although they possess small conserved segments [44,45]. As opposed to miRNAs, which negatively regulate protein-coding genes by targeting the 3′ UTR (untranlated rigions) of their target mRNAs, longer ncRNAs can negatively or positively regulate protein-coding genes [46–48]. Moreover, the known mechanisms underlying the function of lncRNAs are diverse. Considering the dramatic differences in the functions and mechanisms between miRNAs and lncRNAs, distinct evolutionary constraints on these types of RNAs can be easily imagined. Interestingly, a number of researchers have observed the structural conservation of lncRNAs [49]. Some specific structural regions of lncRNAs seem to play regulatory roles, whereas other regions consisting of exact sequences serve only as linkers between different functional modules [50–52].

Various technologies have emerged to detect the stability and conservation of ncRNA. Transcriptome sequencing has been applied to the profiling of ncRNAs conservation. First-generation sequencing, proposed by Sanger [53], was based on double deoxidizing chain-termination. The sequencing reads associated with this method can reach 1000 bp with high accuracy; however, due to its high cost and low throughput, it has not been widely used. Compared with first-generation sequencing, next-generation sequencing (NGS) [54] accelerates the sequencing process and exhibits high throughput. Moreover, NGS can be simultaneously utilized for the analysis of ncRNA expression profiling and sequence variation. Whereas the sequencing reads dramatically decrease, polymerase chain reaction (PCR) leads to a high sequencing error rate. Without PCR amplification, single-molecule sequencing (SMS) decreases the error rate associated with NGS. In addition, SMS maintains a high throughput status, a low cost and the ability to produce long sequencing reads [55]. In addition to directly determining the sequences of ncRNAs, there are several databases and types of software to profile the evolutionary conservation of ncRNAs. Basic Local Alignment Search Tool (BLAST) [56], which is based on sequence alignment, is the most widely used program to search for sequence similarity. Other packages such as CLUSTAL X [57], mirTools [58] and MEGA3 [59] have been used for analyzing data derived from NGS to profile the sequence conservation of ncRNAs. Moreover, numerous emerging structure prediction methods offer a new strategy for the study of the structural conservation and function of ncRNAs.

3.2. Roles of ncRNAs and the Mechanisms Involved in Their Functions

Many ncRNAs remain undiscovered, and the functions of the majority of previously discovered ncRNAs are not yet known. Furthermore, a low evolutionary conservation of these RNAs has been verified. All of these indications suggested that ncRNAs do not possess biological function. However, mounting evidence suggests that the lack of sequence conservation does not necessarily symbolize a deficiency in function [40]. Increasingly, studies have revealed that ncRNAs are involved in gene expression at almost every level of organismal differentiation and development, impacting processes including transcriptional/post transcriptional regulation, chromatin architecture, translation, alternative splicing of pre-mRNA and many other biological processes [15–17,60].

Several mechanisms by which ncRNAs regulate gene expression have been discovered. (1) Piwi-interacting RNAs (piRNAs) have been shown to play roles in the formation of heterochromatin [61]; (2) miRNAs are involved in the modification of histones [62]; (3) various ncRNAs, such as miRNAs and promoter RNAs (pRNAs), play regulatory roles in DNA methylation [63]; (4) some lncRNAs impact the alternative splicing of pre-mRNA [64]; (5) miRNAs degrade mRNAs through binding to the 3′ UTR of their target mRNAs [7]; (6) lncRNAs modulate their structures to recruit specific proteins and form a complex [65]; and (7) interestingly, the newly discovered organellar ncRNAs have been shown to be associated with mitochondria and chloroplasts and represent an emerging mechanism underlying ncRNA-regulated gene expression [66].

3.3. ncRNAs in Diseases and Clinical Diagnosis

Because ncRNAs regulate various levels of gene expression and are involved in numerous biological processes, the dysregulation of ncRNAs is linked to diseases. It has been reported that ncRNAs exert significant effects on the immune response, inflammatory lung diseases [67], neurodevelopmental disorders [68] and cancer [69–71].

In general, abnormal tissues are obtained by invasive methods for the detection of biomarkers in the diagnosis or clinical treatment of tumors. However, due to the introduction of an external source, this is not the optimal choice for diagnostic and therapeutic applications. The characteristics of stability, specificity, sensitivity, predictability and accessibility are required for quantifiable indicators of diseases [72]. Some ncRNAs have been demonstrated to have potential as biomarkers and therapeutic targets for diseases due to their stabilities and accessibilities without invasive obtainment methods [73]. miRNAs are stable and have been found in biological fluids such as urine, serum, saliva and plasma, allowing miRNAs to be easily detected via non-invasive methods [74,75]. The detection of aberrant expression of miRNAs has been applied to the diagnosis and prognosis of cardiac diseases [76] and autoimmune diseases [77]. A genome-wide analysis has revealed only a fraction of lncRNAs are unstable and surprisingly, intronic, intergenic and cis-antisense lncRNAs are highly stable with a half-life of more than 16 h [78]. Some serum-derived lncRNAs have been used as biomarkers for hepatocellular carcinoma and colorectal cancer with high stability, reproducibility and specificity [79]. Moreover, snoRNAs serve as potential biomarkers for the diagnosis of non-small cell lung cancer (NSCLC) [80] and osteoarthritis progression after anterior cruciate ligament (ACL) injury [81]. Undoubtedly, the understanding of ncRNA function contributes to the development of biomarkers for the prognosis and clinical treatment of diseases.

4. Long Noncoding RNAs

Long noncoding RNAs (lncRNAs) consist of at least 200 nucleotides [82]. The structural conservation of lncRNAs is stronger than the conservation of their nucleotide sequences. It has been recognized that lncRNA transcription regulates the expression of genes in close genomic proximity in a cis-acting manner [83–88] and targets distant transcriptional activators or repressors in a trans-acting manner [89,90]. Additionally, various mechanisms involved in the transcriptional regulation of lncRNAs have been elucidated (some examples are shown in Figure 2) [83,84]. Moreover, lncRNAs also participate in epigenetic gene regulation [91,92]. Models of their functions are shown in Figure 2, where lncRNAs are depicted as playing a variety of roles in cellular networks. Therefore, it is inevitable that the dysregulation of lncRNAs is closely associated with diseases [18,19,93].

4.1. Evolutionary Conservation of lncRNAs

Ken C. Pang *et al.* [40] investigated several types of noncoding RNAs that have been demonstrated or predicted to possess functionality, including miRNAs, lncRNAs and snoRNAs. As expected, lncRNAs are less conserved than miRNAs and snoRNAs. However, their findings imply that this lack of conservation does not necessarily dictate a lack of function. Due to the absence of conservation at the nucleotide sequence level, functional studies of lncRNAs are challenging. A number of researchers

have uncovered a structural conservation [49]. Some specific structural regions of lncRNAs seem to play regulatory roles, while other regions consisting of exact sequences serve only as linkers between different functional modules [50–52].
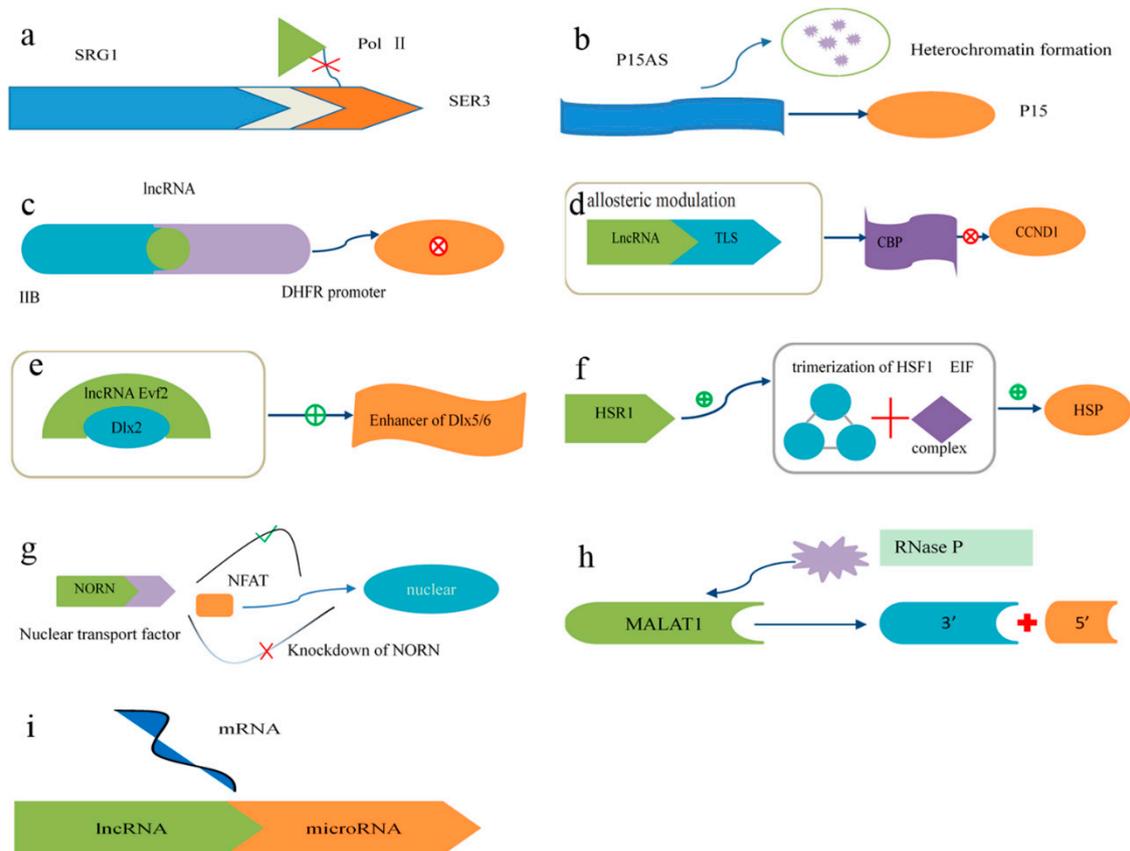
Figure 2. Mechanisms of lncRNA action in transcriptional regulation. (**a**) Transcription of the lncRNA SRG1 inhibits the expression of the SER3 gene by interfering with the binding of RNA polymerase II to DNA; (**b**) The expression of the p15 antisense RNA, the lncRNA of a tumor suppressor gene, results in the silencing of the p15 gene through the induction of heterochromatin formation, which persisted after the p15 antisense RNA was turned off; (**c**) lncRNA binds to the major DHFR promoter and IIB, a general transcriptional factor, to form a stable and specific complex to dissociate the preinitiation complex from the major DHFR promoter; (**d**) As a response to stress, the RNA-binding protein TLS, under allosteric modulation via lncRNA upstream of CCND1, binds to chromatin-binding protein (CBP) and inhibits CBP/P300 HAT activities on CCND1; (**e**) The lncRNA Evf2, a crucial co-enhancer of regulatory proteins involved in transcription, cooperates with the Dlx2 protein to activate the Dlx5/6 enhancer in a target gene; (**f**) In response to heat shock, the lncRNA HSR1 (heat shock RNA-1) promotes the trimerization of HSF1 (heat-shock transcription factor 1), and consequently the translation factor EIF interacts with HSR1 and HSF1 to forms a complex to facilitate the expression of heat-shock protein (HSP); (**g**) NFAT is nuclear factor of activated T cells. The lncRNA NRON (noncoding repressor of NFAT) may form a complex with importin proteins to regulate the subcellular localization of NFAT. The knockdown of NRON increases the expression and activity of NFAT; (**h**) The lncRNA metastasis-associated lung adenocarcinoma transcript 1(MALAT1) has been shown to be abnormally expressed in many human cancers. The nascent MALAT1 transcript is cleaved by RNase P to produce the 3′ end of the mature MALAT1 transcript and the 5′ end of the small RNA; (**i**) Several studies have elucidated that some lncRNAs can act as microRNA sponges to competitively bind to microRNAs and decrease microRNA-induced tumorsphere differentiation.

4.2. Mechanisms of lncRNA Function

The known mechanisms involved in the function of lncRNAs are as follows (Figure 2): (a) To induce transcriptional interference, lncRNAs spanning downstream promoter regions of protein-coding genes interfere with transcription factors via binding to their activators and repress the expression of these protein-coding genes [94]; (b) To initiate chromatin remodeling, the transcription of lncRNAs may induce heterochromatin formation and DNA methylation, thus leading to the silencing of tumor suppressor genes [47,95]; (c) lncRNAs bind to basal transcription factors to inactivate their promoters and thus repress the expression of target genes [96]; (d) lncRNAs activate accessory proteins to repress gene expression [83,97]; (e) lncRNAs activate transcription factors to promote the expression of target genes. This reveals a novel mechanism involving the cooperative actions of an lncRNA and a homeodomain protein to regulate transcription [98]; (f) The formation of a trimer containing an activator protein, a translation elongation factor and an lncRNA accelerates the expression of target genes [99]; (g) lncRNAs interact with importin proteins to regulate the subcellular localization of transcription factors. [100]; (h) lncRNAs act as the precursors of small RNAs to perform functions [101]; (i) lncRNAs bind to small RNAs to modulate their activities [102].

4.3. Epigenetics

It is reported that lncRNAs participate in the epigenetic regulation of gene expression [103–106], and recent studies suggests a unified model of their mechanism of action. The lncRNAs may directly or indirectly recruit protein complexes involved in chromosome modifications, which results in epigenetic regulation [91]. In accordance with the relative positional relationship between lncRNAs and their target genes, mechanisms by which lncRNAs regulate target genes can be considered cis [84–88] or trans [83,89]. For those lncRNAs regulating target genes in cis, it was found that the RNAs can form a nuclear complex that is closely linked to the silenced genes. It is suggested that the lncRNAs may bind to epigenetics modifiers to mediate gene silencing [107]. The HOTAIR lncRNA inactivates genes in trans and interacts with Polycomb Repressive Complex 2 (PRC2) to mediate transcriptional silencing of the HOXD locus [90].

4.4. LncRNAs and Disease

As mentioned above, an increasing number of studies have demonstrated that lncRNAs participate in th As previously mentioned, an increasing number of studies have demonstrated that lncRNAs participate in the regulation of protein coding genes at the transcriptional and posttranscriptional levels [108]. It is reported that the dysregulation of lncRNAs seems to be the primary cause of many complex human disease processes [109,110], including the development and progression of many types of cancer [111], such as colon cancer [112], prostate cancer [113], breast cancer [114], liver cancer [115], gastrointestinal cancer [116] and other cancers [12,117]. Moreover, some studies have shown aberrant lncRNA expression in neurological diseases [118,119]. Further, mounting studies have suggested potential roles for lncRNAs in immunity [120,121].

4.5. lncRNA Structure and Function

Similar to mRNAs, distinct mature ncRNAs can be obtained from primary non-protein coding RNA transcripts via alternative splicing in various differentiated cells, developmental stages or physiological states. It has been estimated that 95% of human primary transcripts of genes containing multiple exons are regulated by alternative splicing [122]. Alternative splicing produces transcript diversification [123]. Alternative splicing of pre-mRNAs generates circular RNA (circRNA) isoforms, ncRNAs with circular structures formed by covalent bonds without a 5′ terminal cap or a poly A tail [124]. In general, canonical splicing processes pre-mRNA sequentially in a 5′ to 3′ direction. The processing involves two transesterification reactions to form the intron lariat, followed by the orderly linkage of upstream and downstream exons [125]. However, in the models of the formation

of circRNAs, the presence of a non-canonical transcription start determines that an orphan upstream 3′ exon splice site could be generated and then paired with a downstream 5′ exon splice site with introns being excised, which produces a circRNA with a circular structure [124]. Trans-splicing and exon skipping are two potential mechanisms by which circRNAs can be generated [126]. Alternative splicing produces many isoforms of the new discovered lncRNA ANRIL associated with different expression patterns and single nucleotide polymorphisms (SNPs). In general, introns are rapidly excised after transcription. However, more than 100 human introns have their 3′ tails degraded but retain their 2′,5′-phosphodiester bond at the splice site without being hydrolyzed. The reserved introns accumulate to form circular intronic lncRNAs (ciRNAs). At the 5′ and 3′ ends of ciRNAs, there are snoRNA structures that replace the 5′ cap and poly A tail and facilitate the accumulation of ciRNAs [127]. Existing evidence has shown that ciRNAs play cis-regulatory roles in the transcription of their parental genes through an interaction with the Pol II machinery [128]. The early discovered lncRNA Nuclear Enriched Abundant Transcript 1 (NEAT1) (MEN ε/β) has been shown to generate distinct isoforms (MEN ε and MEN β) by the alternative processing of the NEAT1 3′ end. MEN ε is characterized by poly A at its 3′ end, whereas, similar to the lncRNA (MALAT1), the 3′ end of MEN β consists of a triple helix structure [129]. Intriguingly, the structure of MEN β is more stable in various species, and the reason for this is currently under investigation [130].

It is currently accepted that the explanation for the various functions of lncRNAs lies in their multiple structures. Mounting evidence has revealed that some lncRNAs and circRNAs can serve as miRNA sponges and inhibit the binding of miRNAs to their target mRNAs to perform their functions [131]. Maternally expressed gene 3 (MEG3), which is highly expressed in the human pituitary, is an imprinted gene that can exist as 12 different transcriptional isoforms due to alternative splicing. All of the MEG3 isoforms have been recognized to inhibit tumor cell growth. The secondary structure motifs M1, M2 and M3 were observed in all of the MEG3 isoforms, and the M2 and M3 motifs have been shown to be closely involved in the activation of P53 and the inhibition of tumor cell growth [132]. However, some lncRNA isoforms perform opposing roles in biological processes. It is reported that the tumor suppressor gene PTEN is regulated by its pseudogene (PTENpg1) through the miRNA sponge action of PTENpg1. To further investigate this regulatory mechanism, two PTENpg1 antisense RNAs (asRNAs) were discovered to play opposing roles in the regulation of PTEN [133]. X-chromosome inactivation (XCI) is a common phenomenon in epigenetic processes. The lncRNA Xist (X-inactive specific transcript) is reported to act as a critical suppressor of X-chromosome inactivation (XCI) [134–136]. Several tandem repeat units composed of two stem-loop structures at the 5′ end of Xist have been shown to be essential for the initiation of XCI [51]. Circular ANRIL (cANRIL) is an ANRIL isoform whose circular structure is a by-product of pre-mRNA alternative splicing. Previous studies suggest that alterations of the structure and/or expression of ANRIL isoforms regulate the expression of INK4/ARF and are associated with atherosclerotic vascular disease (ASVD) [137]. MALAT1, also called nuclear-enriched transcript 2 (NEAT2), has been used as a prognostic marker for the occurrence and development of several types of tumors [138–140]. At the post-transcriptional level, the specific secondary structure at the 3′ end of the MALAT1 primary transcript can be recognized by RNase P and RNase Z, generating a triple helix structure that stabilizes MALAT1 and enables MALAT1 to perform its functions [129,141]. The ncRNA growth arrest-specific 5 (Gas5) is predicted to contain several specific hairpin structures and to be involved in starvation-induced cell survival and metabolic activities through the regulation of glucocorticoid receptor (GR) transcription [142].

5. Structural Prediction of ncRNAs

To elucidate the functions of lncRNAs and to further investigate the question of whether nucleotide sequences serve as functional units or simply linkers of different functional modules, it is necessary to study the structures of lncRNAs and the interaction between their structure and sequence. RNA possesses a unique ability to form complex secondary and tertiary folds [29]. It has been gradually recognized that the structural flexibility of RNA enables it to perform organizational,

catalytic and regulatory functions [25,142,143]. It is now becoming feasible to obtain the functional annotation of transcriptomes based on RNA structure [28]. Traditional methods to investigate RNA structure include chemical probing [144], X-ray crystallography and NMR [145–148]. However, an increasing number of lncRNA molecules have been discovered. Due to the rapid degradation and difficult crystallization of RNA molecules, it is difficult to determine their stereo-chemical structure with these traditional approaches [28]. It is necessary to develop powerful computational methods to predict RNA structure. In this section, various structure prediction methods for noncoding RNAs are reviewed.

5.1. Prediction of ncRNA Secondary Structure

The folding process of the majority of RNA molecules represents a transition from secondary to tertiary structure [149]. Therefore, obtaining the RNA secondary fold is the first step in exploring the functions of ncRNAs [29]. In recent years, various methods have been proposed for predicting RNA secondary structure. These methods are based on two distinct ideas: multiple sequence alignments and the minimum free energy model [28].

5.1.1. Multiple Sequence Alignments

Methods based on comparative sequence analysis rely on the fact that the structural conservation is greater than the sequence conservation of RNA [150,151]. Comparative sequence analysis compares several RNA sequences with similar secondary structures to search for conserved secondary structural units and predicts the secondary structure of an unknown RNA sequence [152].

Foldalign

Foldalign [30], simplified from Sankoff [153], utilizes a dynamic programming algorithm to find the highest scoring local alignment between a sequence and an alignment of other sequences or between two sequences [154]. The correlation coefficient [155] ranges from 0.8 to 0.9 between the verified database and the predicted structural alignments. Foldalign compares each sequence with every other sequence, and the numbers of the highest scoring alignments are saved. It can effectively perform on RNA sequences less than 300 nt. In addition, the time associated with this method is significantly reduced compared with the Sankoff version and other variants. However, the speed and efficiency of Foldalign require improvements [154]. The web server can be accessed at http://rth.dk/resources/foldalign/ [156].

Dynalign

Dynalign [157], which is based on a dynamic programming proposed by Sankoff, searches for a structure with low free energy common to two sequences without sequence identity by combining comparative sequence analysis and free energy minimization. Compared with free energy minimization alone, the average accuracy of this algorithm is improved from 47.8% to 86.4% for 5S rRNAs. It can predict a set of suboptimal secondary structures and create dot plots to read the information contained in suboptimal structures. Moreover, enzymatic cleavage data [158] and chemical modification probing experiments [159] can be applied to increase the prediction accuracy. However, it cannot predict pseudoknots, and the calculation is limited to sequences whose lengths are less than 400 nt [160].

Pfold

Pfold [161] is based on the KH-99 algorithm [162], which combined evolutionary information and a probabilistic structure model. Pfold can accommodate larger numbers of sequences, which can compensate for the limitations of the KH-99 algorithm. Due to its high computational speed and prediction accuracy, it is able to predict RNA secondary structure when long sequences and large

numbers of homologous sequences need to be analyzed. With six sequences, an accuracy of 75% is attainable. In addition, many more sequences can be accommodated by Pfold, allowing for even higher accuracies [31]. However, there is still much room for this method to be improved, such as the introduction of a grammar to describe native-like RNA structures, stacking interactions and other models for base-pair evolution [161]. In addition, it cannot predict pseudoknots. Pfold is available through the web-based server www.daimi.au.dk/~compbio/pfold [163].

Alifold

The Alifold service [164,165], an extension of Zuker's algorithm [166], uses modified dynamic programming algorithms combined with a covariance term to compute the consensus secondary structure of a set of aligned RNA sequences. It can predict minimum free energy structures and pair probabilities. The current limit for the length of the alignment is 3000 nt [165]. The advantages and limitations of Alifold are almost identical to those of RNAfold. This service can be accessed via the Vienna RNA web server at http://rna.tbi.univie.ac.at/cgi-bin/RNAalifold.cgi [167].

MARNA

MARNA [168], a non-probabilistic approach [169], performs pairwise alignments considering both the primary and secondary structures. It folds sequences using the minimum free energy and then provides structural alignment among a set of homologous sequences. When the conservative sequence regions are invisible, MARNA is an appropriate option to predict RNA secondary structure. Users can designate individual parameters that can set the weight for either sequence or structural properties. However, the total length of sequences should not be longer than 10,000 nt. MARNA can be used online on the following webpage: http://rna.informatik.uni-freiburg.de/MARNA/Input.jsp [170].

A large number of studies and experiments have demonstrated that comparative sequence analysis processes higher prediction probabilities when the RNA sequence templates have high similarity [171]. However, because comparative sequence analysis depends on the prior knowledge of sequences, this model is unfit for single RNA sequences or sequences from considerably different sources [152]. In addition, comparative sequence analysis is time- and internal storage-consuming, which limits its application for predicting longer RNA sequences [28].

5.1.2. Minimum Free Energy Model

When no prior knowledge is available and only a single sequence is offered, an accurate and popular method is to search the minimum free energy model through thermodynamic computation [172]. This model utilizes efficient dynamic programming algorithms to search for a secondary structure with the minimum free energy [166]. However, true RNA secondary structure may not be the structure with the minimum free energy. Zuker *et al.* [173] developed the concept of suboptimum structures. All suboptimum structures must be further identified by biology researchers.

Mfold

Mfold [32] divides RNA secondary structures motifs into the stem area, bulge loop, internal loop and hairpin loop. Different computational methods are used to calculate the free energy of different motifs. Then, the motifs are assembled through dynamic programming algorithms, and the secondary structure with the minimum free energy can be obtained. Using this method, prior knowledge can be specified before the prediction; the structure of circular RNA sequences is predictable, the maximum for internal or bulge loops can be set, and the maximum distance between paired bases can be artificially determined. Many studies have proposed that RNA secondary structure affects splicing activity [174]. Yun Yang *et al.* [175] discovered that the inherent intronic elements are underlying mechanisms for the pre-mRNA splicing process. These elements have been found conserved at the RNA secondary structural level. In their studies, the Mfold program was used to predict intronic pairings. However,

Mfold can only predict the secondary structure of single stranded RNA. The portal for the Mfold webserver is http://unafold.rna.albany.edu/?q=mfold [176].

RNAfold

RNAfold [33], which is based on dynamic programming algorithms and computations of the equilibrium partition functions and base pairing probabilities, uses the minimum free energy model and multiple sequence alignments when given single stranded RNA sequences and several stranded RNA sequences, respectively. RNAfold is a reliable option regardless of whether the base pairing of G and U is acceptable or not. Moreover, the sequences can contain incorrect characters. Furthermore, the program can predict single stranded and several stranded RNAs. Humann *et al.* [177] discovered differentially expressed lncRNAs in the larval ovaries of honeybee caste by using the RNAfold program and other biological technologies. They named the newly discovered lncRNAs lncov1 and lncov2. The secondary structures of both RNAs consist of several consensus hairpin motifs lacking coding potential. However, it is worth noting that the length of the sequence should not be more than 300 nt. When predicting several stranded RNAs, the program can only produce the consensus structure as opposed to the secondary structure of each sequence. In addition, the total length of the sequences cannot exceed 10K nt when predicting the consensus structure. The portal for the RNAfold web server is http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi [178].

RNAshapes

RNAshapes [34], based on the abstract shapes approach [179], is a new method that combines three RNA analysis tools: the analysis of shape representatives, the consensus shapes approach and the calculation of shape probabilities. Compared with other current RNA folding algorithms, RNAshapes only describes classes of structures from concrete secondary structures. These structures fall into different shape categories. Within a shape class, every representative is the secondary structure with the minimum free energy. Using this package, the single-stranded RNA, the sequence files and the multi-sequence files are all predictable. For a given threshold value, the number of shapes is less than the number of structures, and the native structures are among the shape representatives. Therefore, users can avoid researching redundant suboptimal structures [179]. However, because the folding kinetics are not considered, the minimum free energy prediction may be incorrect. RNAshapes is freely available at http://bibiserv.techfak.uni-bielefeld.de/rnashapes [180].

RNAstructure

RNAstructure [35] utilizes the most recent set of thermodynamic parameters to implement the nearest neighbor parameters as determined by the Tuner group [181,182] based on dynamic programming algorithms and Sankoff, which allow sequence alignment and structure prediction to proceed simultaneously. The user interface is friendly and powerful. Its "Max % Energy Difference" and "Max Number of Structures" can be modified to limit the number of suboptimal structures predicted. Moreover, experimental data can be added to constrain the structures. Furthermore, it can predict both single stranded RNA and a structure common to two sequences. This method has been widely used in research. Ding *et al.* [183] compared the structural features of mRNAs *in vivo* with predicted structures (determined by RNAstructure) *in silico* and revealed that mRNAs related to stress responses have structural features, such as longer maximal loop length and more single strandedness, that allow for easy conformational changes under various environmental conditions. SPRY4-IT1, the lncRNA that regulates invasion and apoptosis, was predicted to contain long hairpin motifs (by RNAstructure), suggesting that SPRY4-IT1 may function as an RNA molecule [184]. The package is available for downloading at http://rna.urmc.rochester.edu/RNAstructure.html [185].

The information regarding the methods described above is summarized in Table 1. Apart from the mainstream methods mentioned above, Sfold [186,187], Contrafold [188], and MPGAfold [187] are also available to solve problems when predicting RNA secondary structure. Although there has been

remarkable development in the methods to predict RNA secondary structures, the methods based on the free energy parameters proposed by Zuker *et al.* [32,173] still represent the mainstream.

Table 1. Comparison of the various major methods to predict RNA secondary structure.

Method	Principle	Advantages	Limitations
Foldalign [30]	Sankoff, dynamic programming algorithm	time complexity decreased	length of sequence shorter than 300 nt; low speed and efficiency
Dynalign [158]	Sankoff, dynamic programming algorithm	suboptimal secondary structures accessible; constrained information added	Pseudoknots not predictable
MARNA [169]	folding sequences using minimum free energy; proceedings structural alignment	individual parameters freely set	total length of sequences shorter than 10,000 nt
Mfold [32]	Zuker's dynamic programming algorithm based on minimum free energy model	priori knowledge specified; structure of circular RNA sequence predictable; some values related to structure artificially made	only structure of single stranded RNA predictable
Alifold/RNAfold [33,165]	minimum free energy model and multiple sequence alignment	containing incorrect characters; single stranded RNA and several stranded RNAs predictable; base pairing of G and U acceptable	when predicting several stranded RNAs, only producing consensus structure instead of the secondary structure of each sequence; when predicting single sequence, its length requirement is less than 300 nt; total length of sequence not to exceed 10K nt when predicting consensus structure
RNAshapes [34]	abstract shapes approach	single stranded RNA, sequence files and multi-sequence files predictable; redundant suboptimal structures avoided	does not consider folding kinetics; minimum free energy prediction may be incorrect
RNAstructure [35]	dynamic programming algorithm and Sankoff	number of suboptimal structures limited; structures constrained by experimental data	only AGCU predictable

5.2. Prediction of ncRNA Tertiary Structure

The formation of specific tertiary structures is essential for the functioning of noncoding RNAs in many biological processes [189]. RNAs can alter their tertiary structure under different conditions, enabling them to interact with other RNAs, ligands, proteins or themselves [28]. In this section, methods to predict the tertiary structure of ncRNAs are reviewed.

FARNA

FARNA [190], derived from the Rosetta methods of protein tertiary structure prediction [191], utilizes coarse-grained models as dummy atoms to replace the center of each base and seek RNA tertiary structure with the minimum free energy. The prediction accuracy of the main chains can reach a 4 Å root-mean-square-deviation (RMSD) [192] for short RNA sequences with a length less than 30 nt. The prediction accuracy of this method can be further improved by combining it with experimentally determined secondary structure information [193]. In recent years, Baker *et al.* [194] have introduced all-atom items to FARNA, which has allowed FARNA to become an all-atom structure prediction method. FARNA is characterized by a better computational efficiency in comparison with numerous

sampling strategies. However, FARNA can only predict the tertiary structure of small RNA molecules (<40 nt). Challenges remain in accommodating RNA molecules of longer lengths or with complex topological structures.

NAST

NAST (The Nucleic Acid Simulation Tool) [195], based on coarse-grained models, uses knowledge-based energy functions to automatically predict RNA tertiary structure. NAST requires secondary and tertiary contact information for target RNA molecules to direct folding. It has a mean RMSD of 8.0 ± 0.3 and 16.3 ± 1.0 Å for the yeast phenylalanine tRNA and the P4–P6 domain of the Tetrahymena thermophila group I intron, respectively. Plausible RNA structures can be created with empirical RNA geometric distributions, a relatively high modeling speed can be achieved by using single-point-per-base models, and the capacity to constrain and filter models with experimental data improves the prediction accuracy of NAST. Due to computational complexity, modeling large RNA molecules remains difficult. The software package is freely available at https://simtk.org/home/nast [196].

iFoldRNA

iFoldRNA [37] uses discrete molecular dynamics (DMD) to rapidly explore RNA tertiary conformation [36,197]. Compared with traditional dynamic molecule simulations, the rapid conformation sampling ability of DMD contributes to its rapid structure prediction [198]. Low RMSDs (2–3 Å) are observed in the predictions of iFoldRNA. iFoldRNA can predict the tertiary structure of small RNA molecules (<50 nt) with simple topological structure. When predicting larger RNA molecules (>50 nt), a longer time is required to sample conformational space, which exponentially increases. Recently, parameters including base pairing, base-stacking, and hydrophobic interactions obtained from experiments have been integrated into iFoldRNA to constrain the structures of larger RNA molecules [199].

BARNACLE

BARNACLE [200], a probabilistic model of RNA structure, provides sampling of RNA conformations in continuous space. The current state of prediction methods such as FARNA are primarily based on combining short fragments obtained from experiments to construct reasonable native-like tertiary structures. However, there are some computational sampling problems associated with these methods. It is possible for BARNACLE to efficiently sample 3D conformations of RNA on a short length scale. BARNACLE can accurately predict RNA tertiary structure when the length of the RNA sequence is less than 50 nt (10 Å RMSD). Nevertheless, structure sampling becomes difficult due to too many degrees of freedom with longer RNA molecules or with those that harbor complicated topological structures. Moreover, the sequence and evolutionary information of BARNACLE needs to be extended.

CG Model

The CG model [201] models RNA structures with molecular dynamics based on a new statistical coarse-grained potential. The statistical analysis of 688 RNA experimental structures has been applied to parameterize the CG potential [202]. The computational efficiency is greater than that of the all-atom model because of the reduction in the number of angles, bonds and torsion calculations. Fifteen RNA molecules with a length of 12 to 27 nt have been tested through molecule dynamics simulation, this shows that 75% of RNA molecules can be led to native-like structures with at least one out of multiple pathways using the simulated annealing method. If secondary or tertiary structure interaction information is provided, all of the RNA molecules will successfully be folded into structures with an RMSD less than 6.5 Å. Similar to other methods, this method is restricted to predicting small RNA molecules with simple topological structures.

RNA2D3D

RNA2D3D [203], different from other structure prediction methods, is based on unpaired bases derived from Assisted Model Building with Energy Refinement (AMBER) [204] and canonical base-pairings of the A-form helix to model RNA tertiary structure. However, overlapping atoms, covalent bond disassociation and other structural problems that exist in the RNA tertiary structure are automatically generated by RNA2D3D. Therefore, further optimization is necessary to obtain a reasonable RNA tertiary structure. After the adjustment and optimization of RNA2D3D, the pseudoknot structure of the telomerase RNA, with a length of 48 nt, has been successfully built by Shapiro *et al.*, and the RMSD reached 7 Å [205–207].

Vfold Model

The Vfold model [208] is a physics-based method for predicting larger and more complex RNA molecules from nucleotide sequences. This method uses a multi-scaling strategy in which secondary and tertiary structures are obtained in a serial fashion. Compared with other methods, the Vfold model can predict larger RNA molecules, for example the 122-nt 5S rRNA domain (RMSD 7.4 Å). The most significant advantage of the Vfold model is its statistical mechanical calculations for the conformational entropy of RNA tertiary structures. In addition, the model can be used to predict all low-lying tertiary structures in the energy landscape. However, this method does not consider the sequence-dependent tertiary contacts, such as general loop-loop and loop-helix interactions, in loop-free energy minimization.

RSIM

RSIM [36], a fully automated application, is an improved approach to predict RNA tertiary structure using the fragment assembly method based on RNA secondary structure constraints. It overcomes the pitfalls of FARNA, such as the reduction of the size of the sampled conformational space and the reasonable base-pairing constraint using the fragment assembly method. Monte Carlo simulations, a statistical potential and a diverse fragment library are further used to refine the tertiary structures obtained by RSIM. During the refinement, the stimulation paths can be tracked. RSIM can accommodate RNA molecules with a length over 40 nt (RMSD 4.8 Å). However, RSIM cannot automatically predict the tertiary structure of RNA molecules with pseudoknot structures. RSIM is available at http://www.github.com/ jpbida/rsim [209].

3dRNA

3dRNA [38], based on RNA sequence and secondary structural information, is a method for the rapid and automated building of RNA tertiary structure. It is a hierarchical approach to the construction of RNA tertiary structure [210]. Compared with other methods, 3dRNA can obtain RNA tertiary structural templates from different RNA families. It is found that the conformations of the backbone of RNA structural templates of the same sequence are similar to each other. These changes contribute to a high average prediction accuracy of 3.97 Å RMSD. 3dRNA is not limited to predicting the tertiary structures of small RNA molecules or those with simple topology. For RNA molecules of a large size and complex topology, the predicted tertiary structures have an average RMSD of 5.7 Å. The research conducted in Qian's lab in Northwestern Polytechnical University has predicted the tertiary structures of 5 lncRNAs with 3dRNA and uncovered important roles for these lncRNAs in bone formation when MACF1 (Microtubule actin cross-linking factor l) is down-regulated (data not shown). The package is available at http://biophy.hust.edu.cn/3dRNA/3dRNA-1.0.html [211].

The methods mentioned above are widely used to predict RNA tertiary structure. Furthermore, MC-Fold/MC-Sym [212], based on the nucleotide cyclic motif (NCM), is a first-order object to represent nucleotide relationships in structured RNAs. ASSEMBL [213] is an interactive graphical tool based on human-computer interactions to analyze and build 2D and 3D RNA models. In general, the prediction

accuracy of RNA tertiary structure will be largely improved by the addition of structural information, such as RNA secondary structure, distance, rotation angle, dihedral angle and other tertiary structural information [214]. However, Liang and Schlick [215,216] accessed these existing RNA tertiary structure prediction methods and found that they are restricted to analyzing short (<50 nt) or topologically simple molecules with RMSD less than 6 Å. When predicting larger (50 to 130 nt) or more topologically complex RNA molecules, the tertiary structure can be obtained with a mean RMSD of 20 Å. Moreover, the existing prediction methods for RNA tertiary structure require human-computer interactions for further adjustment to optimize the obtained RNA tertiary structure. Therefore, the proposal of 3dRNA is a significant step forward in the prediction of RNA tertiary structure. The various methods for predicting RNA tertiary structure are summarized in Table 2.

Table 2. Various methods for predicting RNA tertiary structure.

Method	Principles	Advantages	Limitations
FARNA [36]	coarse-grained models, minimum free energy	better computational efficiency	small RNA molecules (<40 nt)
NAST [195]	coarse-grained models, knowledge-based energy function	relatively high modeling speed; constraint models	computational complexity
iFoldRNA [37]	discrete molecular dynamics	rapid conformational sampling ability	small RNA molecules (<50 nt)
BARNACLE [201]	probabilistic model, sampling of RNA conformations in continuous space	efficient sampling of 3D RNA conformations on a short length scale	small RNA molecules (<50 nt); sample difficulty
CG Model [202]	molecular dynamics based on a new statistical coarse-grained potential	high computational efficiency	small RNA molecules or those with simple topology
RNA2D3D [204]	base-pairing structure of RNA molecules	can predict pseudoknots	obtaining reasonable RNA tertiary structure to be solved
Vfold Model [209]	physics-based method	statistical mechanical calculations for the conformational entropy of RNA tertiary structures	does not consider the sequence-dependent tertiary contacts
RSIM [36]	fragment assembly	the reduction in the size of conformational space sampled; reasonable base-pairing constraints	RNA molecules with pseudoknot structures not automatically predictable
3dRNA [38]	hierarchical approach to construct RNA tertiary structure	highest prediction accuracy	ability to model larger RNA molecules or those with complex topology

6. Conclusions

With an increasing number of studies focused on lncRNAs, an increased understanding of lncRNAs has been achieved. lncRNAs play biological roles in organisms, and their dysregulation is strongly linked to the occurrence and development of various diseases [217]. However, the in-depth knowledge of the function of lncRNAs is a developing but difficult field due to the diversity and complexity of the mechanisms underlying lncRNAs. As RNA function is closely associated with its structure [24], analyzing RNA structure provides a new approach to the study of lncRNAs [28]. Before major progress in the determination of ncRNA structure using physical methods is achieved, the structural prediction of ncRNAs will be a hotly debated issue. At present, the prediction of pseudoknots is very difficult [218]. Our knowledge of thermodynamics [182,219] and algorithms to model RNA molecules undergoing conformational changes [28] is incomplete. These represent problems that need to be addressed for the secondary structure prediction of ncRNAs. Moreover, with the ongoing improvements in the accuracy of ncRNA structural prediction, it is possible to reliably

predict the tertiary structure of small RNA molecules; however, predicting the structure of large RNA molecules or those with complex topological structures [38] remains challenging. Moreover, tackling the structure of non-canonical base pairings in the prediction of RNA tertiary structure remains a difficult problem [38]. Furthermore, to elucidate the complicated mechanisms of actions of lncRNAs, the use of experimental data as constraint information is inevitable. It is expected that issues occurring in the structural prediction of lncRNAs will be addressed in the future and that additional techniques will be applied to studies of lncRNA function, which will allow the further analysis of their functions, molecular regulation and pathological mechanisms in diseases. In the future, lncRNAs may serve as drug targets and provide new opportunities for the treatment of diseases.

Author Contributions: Kun Yan wrote the whole manuscript. Airong Qian and Tuanming Yang made the outline of the manuscript. Fan Zhao, Zhihao Chen, Chong Yin searched for literatures cited in introduction, prediction of RNA secondary structure and tertiary structure respectively. Yasir Arfat, Dijie Li made corrections for the manuscript. Yulong Sun and Lifang Hu changed the format of the manuscript according to the instructions for authors. All of the authors proofread and approved the final manuscript.

Conflicts of Interests: The authors declare no conflict of interest.

Abbreviations

lncRNAs: long noncoding RNAs; ncRNAs: noncoding RNAs; NATs: natural antisense transcripts; PRC2: Polycomb Repressive Complex 2; NAST: The Nucleic Acid Simulation Tool; AMBER: Assisted Model Building with Energy Refinement; MACF1: Microtubule actin cross-linking factor l.

References

1. Mattick, J.S. Non-coding RNA. *Hum. Mol. Genet.* **2006**, *15*, 17–29. [CrossRef] [PubMed]
2. The ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **2007**, *447*, 799–816.
3. Consortium, T.F.; Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; *et al*. The Transcriptional Landscape of the Mammalian Genome. *Science* **2005**, *309*, 1559–1563. [CrossRef] [PubMed]
4. Zhang, Y.; Yang, L.; Chen, L.L. Life without A tail: New formats of long noncoding RNAs. *Int. J. Biochem. Cell Biol.* **2014**, *54*, 338–349. [CrossRef] [PubMed]
5. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; *et al*. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338. [CrossRef] [PubMed]
6. Furuno, M.; Pang, K.C.; Ninomiya, N.; Fukuda, S.; Frith, M.C.; Bult, C.; Kai, C.; Kawai, J.; Carninci, P.; Hayashizaki, Y.; *et al*. Clusters of internally primed transcripts reveal novel long noncoding RNAs. *PLoS Genet.* **2006**, *2*, e37. [CrossRef] [PubMed]
7. Carthew, R.W.; Sontheimer, E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655. [CrossRef] [PubMed]
8. Yang, L.; Duff, M.O.; Graveley, B.R.; Carmichael, G.G.; Chen, L.L. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **2011**, *12*, R16. [CrossRef] [PubMed]
9. Ponjavic, J.; Ponting, C.P.; Lunter, G. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res.* **2007**, *17*, 556–565. [CrossRef] [PubMed]
10. Pefanis, E.; Wang, J.; Rothschild, G.; Lim, J.; Chao, J.; Rabadan, R.; Economides, A.N.; Basu, U. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* **2014**, *514*, 389–393. [CrossRef] [PubMed]
11. Yan, W. Potential roles of noncoding RNAs in environmental epigenetic transgenerational inheritance. *Mol. Cell. Endocrinol.* **2014**, *398*, 24–30. [CrossRef] [PubMed]

12. Batista, P.J.; Chang, H.Y. Long noncoding RNAs: Cellular address codes in development and disease. *Cell* **2013**, *152*, 1298–1307. [CrossRef] [PubMed]

13. Booton, R.; Lindsay, M.A. Emerging role of MicroRNAs and long noncoding RNAs in respiratory disease. *Chest* **2014**, *146*, 193–204. [CrossRef] [PubMed]

14. Wang, P.; Xue, Y.; Han, Y.; Lin, L.; Wu, C.; Xu, S.; Jiang, Z.; Xu, J.; Liu, Q.; Cao, X. The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. *Science* **2014**, *344*, 310–313. [CrossRef] [PubMed]

15. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672. [CrossRef] [PubMed]

16. Guttman, M.; Donaghey, J.; Carey, B.W.; Garber, M.; Grenier, J.K.; Munson, G.; Young, G.; Lucas, A.B.; Ach, R.; Bruhn, L.; *et al.* lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **2011**, *477*, 295–300. [CrossRef] [PubMed]

17. Mercer, T.R.; Dinger, M.E.; Mattick, J.S. Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* **2009**, *10*, 155–159. [CrossRef] [PubMed]

18. Han, P.; Li, W.; Lin, C.H.; Yang, J.; Shang, C.; Nurnberg, S.T.; Jin, K.K.; Xu, W.; Lin, C.Y.; Lin, C.J.; *et al.* A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* **2014**, *514*, 102–106. [CrossRef] [PubMed]

19. Trimarchi, T.; Bilal, E.; Ntziachristos, P.; Fabbri, G.; Dalla-Favera, R.; Tsirigos, A.; Aifantis, I. Genome-wide Mapping and Characterization of Notch-Regulated Long Noncoding RNAs in Acute Leukemia. *Cell* **2014**, *158*, 593–606. [CrossRef] [PubMed]

20. Hindorff, L.A.; Sethupathy, P.; Junkins, H.A.; Ramos, E.M.; Mehta, J.P.; Collins, F.S.; Manolio, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9362–9367. [CrossRef] [PubMed]

21. Chen, G.; Wang, Z.; Wang, D.; Qiu, C.; Liu, M.; Chen, X.; Zhang, Q.; Yan, G.; Cui, Q. LncRNADisease: A database for long-non-coding RNA-associated diseases. *Nucleic Acids Res.* **2013**, *41*, D983–D986. [CrossRef] [PubMed]

22. Sánchez, Y.; Huarte, M. Long Non-Coding RNAs: Challenges for Diagnosis and Therapies. *Nucleic Acid Ther.* **2013**, *23*, 15–20. [CrossRef] [PubMed]

23. Warf, M.B.; Berglund, J.A. The role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem. Sci.* **2010**, *35*, 169–178. [CrossRef] [PubMed]

24. Sim, A.Y.L.; Levitt, M. Clustering to identify RNA conformations constrained by secondary structure. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3590–3595. [CrossRef] [PubMed]

25. Roberts, T.C.; Morris, K.V.; Weinberg, M.S. Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics* **2014**, *9*, 13–20. [CrossRef] [PubMed]

26. Kozak, M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **2005**, *361*, 13–37. [CrossRef] [PubMed]

27. Martin, K.C.; Ephrussi, A. mRNA Localization: Gene Expression in the Spatial Dimension. *Cell* **2009**, *136*, 719. [CrossRef] [PubMed]

28. Wan, Y.; Kertesz, M.; Spitale, R.C.; Segal, E.; Chang, H.Y. Understanding the transcriptome through RNA structure. *Nat. Rev. Genet.* **2011**, *12*, 641–655. [CrossRef] [PubMed]

29. Novikova, I.V.; Hennelly, S.P.; Sanbonmatsu, K.Y. Tackling structures of long noncoding RNAs. *Int. J. Mol. Sci.* **2013**, *14*, 23672–23684. [CrossRef] [PubMed]

30. Havgaard, J.H.; Lyngsø, R.B.; Gorodkin, J. The foldalign web server for pairwise structural RNA alignment and mutual motif search. *Nucleic Acids Res.* **2005**, *33*, W650–W653. [CrossRef] [PubMed]

31. Knudsen, B. Pfold: RNA secondary structure prediction using stochastic context-free grammars. *Nucleic Acids Res.* **2003**, *31*, 3423–3428. [CrossRef] [PubMed]

32. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415. [CrossRef] [PubMed]

33. Hofacker, I.L.; Stadler, P.F. Memory efficient folding algorithms for circular RNA secondary structures. *Bioinformatics* **2006**, *22*, 1172–1176. [CrossRef] [PubMed]

34. Steffen, P.; Voss, B.; Rehmsmeier, M.; Reeder, J.; Giegerich, R. RNAshapes: An integrated RNA analysis package based on abstract shapes. *Bioinformatics* **2006**, *22*, 500–503. [CrossRef] [PubMed]

35. Reuter, J.S.; Mathews, D.H. RNAstructure: Software for RNA secondary structure prediction and analysis. *BMC Bioinform.* **2010**, *11*, 129. [CrossRef] [PubMed]

36. Bida, J.P.; Maher, L.J. Improved prediction of RNA tertiary structure with insights into native state dynamics. *RNA* **2012**, *18*, 385–393. [CrossRef] [PubMed]

37. Sharma, S.; Ding, F.; Dokholyan, N.V. iFoldRNA: Three-dimensional RNA structure prediction and folding. *Bioinformatics* **2008**, *24*, 1951–1952. [CrossRef] [PubMed]

38. Zhao, Y.; Huang, Y.; Gong, Z.; Wang, Y.; Man, J.; Xiao, Y. Automated and fast building of three-dimensional RNA structures. *Sci. Rep.* **2012**, *2*, 734. [CrossRef] [PubMed]

39. Işın, M.; Uysaler, E.; Özgür, E.; Köseoğlu, H.; Şanlı, Ö.; Yücel, Ö.B.; Gezer, U.; Dalay, N. Exosomal lncRNA-p21 levels may help to distinguish prostate cancer from benign disease. *Front. Genet.* **2015**, *6*, 168. [CrossRef] [PubMed]

40. Pang, K.C.; Frith, M.C.; Mattick, J.S. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet.* **2006**, *22*, 1–5. [CrossRef] [PubMed]

41. Lau, N.C.; Lim, L.P.; Weinstein, E.G.; Bartel, D.P. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans. *Science* **2001**, *294*, 858–862. [CrossRef] [PubMed]

42. Lee, R.C.; Ambros, V. An Extensive Class of Small RNAs in Caenorhabditis elegans. *Science* **2001**, *294*, 862–864. [CrossRef] [PubMed]

43. Ambros, V.; Bartel, B.; Bartel, D.P.; Burge, C.B.; Carrington, J.C.; Chen, X.; Dreyfuss, G.; Eddy, S.R.; Griffiths-Jones, S.; Marshall, M.; *et al.* A uniform system for microRNA annotation. *RNA* **2003**, *9*, 277–279. [CrossRef] [PubMed]

44. Nesterova, T.B.; Slobodyanyuk, S.Y.; Elisaphenko, E.A.; Shevchenko, A.I.; Johnston, C.; Pavlova, M.E.; Rogozin, I.B.; Kolesnikov, N.N.; Brockdorff, N.; Zakian, S.M. Characterization of the Genomic *Xist* Locus in Rodents Reveals Conservation of Overall Gene Structure and Tandem Repeats but Rapid Evolution of Unique Sequence. *Genome Res.* **2001**, *11*, 833–849. [CrossRef] [PubMed]

45. Chureau, C.; Prissette, M.; Bourdet, A.; Barbe, V.; Cattolico, L.; Jones, L.; Eggen, A.; Avner, P.; Duret, L. Comparative Sequence Analysis of the X-Inactivation Center Region in Mouse, Human, and Bovine. *Genome Res.* **2002**, *12*, 894–908. [PubMed]

46. Kretz, M.; Siprashvili, Z.; Chu, C.; Webster, D.E.; Zehnder, A.; Qu, K.; Lee, C.S.; Flockhart, R.J.; Groff, A.F.; Chow, J.; *et al.* Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **2013**, *493*, 231–235. [CrossRef] [PubMed]

47. Yu, W.; Gius, D.; Onyango, P.; Muldoon-Jacobs, K.; Karp, J.; Feinberg, A.P.; Cui, H. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* **2008**, *451*, 202–206. [CrossRef] [PubMed]

48. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long noncoding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076. [CrossRef] [PubMed]

49. Smith, M.A.; Gesell, T.; Stadler, P.F.; Mattick, J.S. Widespread purifying selection on RNA structure in mammals. *Nucleic Acids Res.* **2013**, *41*, 8220–8236. [CrossRef] [PubMed]

50. Jeon, Y.; Lee, J.T. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* **2011**, *146*, 119–133. [CrossRef] [PubMed]

51. Wutz, A.; Rasmussen, T.P.; Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **2002**, *30*, 167–174. [CrossRef] [PubMed]

52. Guttman, M.; Rinn, J.L. Modular regulatory principles of large non–coding RNAs. *Nature* **2012**, *482*, 339–346. [CrossRef] [PubMed]

53. Sanger, F.; Nicklen, S.; Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5463–5467. [CrossRef] [PubMed]

54. Niedringhaus, T.P.; Milanova, D.; Kerby, M.B.; Snyder, M.P.; Barron, A.E. Landscape of next-generation sequencing technologies. *Anal. Chem.* **2011**, *83*, 4327–4341. [CrossRef] [PubMed]

55. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B.; *et al.* Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* **2009**, *323*, 133–138. [CrossRef] [PubMed]

56. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [CrossRef] [PubMed]

57. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882. [CrossRef] [PubMed]

58. Wu, J.; Liu, Q.; Wang, X.; Zheng, J.; Wang, T.; You, M.; Sheng Sun, Z.; Shi, Q. MirTools 2.0 for non-coding RNA discovery, profiling, and functional annotation based on high-throughput sequencing. *RNA Biol.* **2013**, *10*, 1087–1092. [CrossRef] [PubMed]

59. Kumar, S.; Tamura, K.; Nei, M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* **2004**, *5*, 150–163. [CrossRef] [PubMed]

60. Amaral, P.; Mattick, J. Noncoding RNA in development. *Mamm. Genome* **2008**, *19*, 454–492. [CrossRef] [PubMed]

61. Brower-Toland, B.; Findley, S.D.; Jiang, L.; Liu, L.; Yin, H.; Dus, M.; Zhou, P.; Elgin, S.C.; Lin, H. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev.* **2007**, *21*, 2300–2311. [CrossRef] [PubMed]

62. Younger, S.T.; Corey, D.R. Transcriptional gene silencing in mammalian cells by miRNA mimics that target gene promoters. *Nucleic Acids. Res.* **2011**, *39*, 5682–5691. [CrossRef] [PubMed]

63. Schmitz, K.M.; Mayer, C.; Postepska, A.; Grummt, I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* **2010**, *24*, 2264–2269. [CrossRef] [PubMed]

64. Beltran, M.; Puig, I.; Peña, C.; García, J.M.; Álvarez, A.B.; Peña, R.; Bonilla, F.; de Herreros, A.G. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial–mesenchymal transition. *Genes Dev.* **2008**, *22*, 756–769. [CrossRef] [PubMed]

65. Mao, Y.S.; Sunwoo, H.; Zhang, B.; Spector, D.L. Direct Visualization of the Co-transcriptional Assembly of a Nuclear Body by Noncoding RNAs. *Nat. Cell Biol.* **2011**, *13*, 95–101. [CrossRef] [PubMed]

66. Dietrich, A.; Wallet, C.; Iqbal, R.K.; Gualberto, J.M.; Lotfi, F. Organellar non-coding RNAs: Emerging regulation mechanisms. *Biochimie* **2015**, *117*, 48–62. [CrossRef] [PubMed]

67. Xie, N.; Liu, G. ncRNA-regulated immune response and its role in inflammatory lung diseases. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **2015**, *309*, L1076–L1087. [CrossRef] [PubMed]

68. Hecht, P.M.; Ballesteros-Yanez, I.; Grepo, N.; Knowles, J.A.; Campbell, D.B. Noncoding RNA in the transcriptional landscape of human neural progenitor cell differentiation. *Front. Neurosci.* **2015**, *9*, 392. [CrossRef] [PubMed]

69. Venkatesh, T.; Suresh, P.S.; Tsutsumi, R. Non-coding RNAs: Functions and applications in endocrine-related cancer. *Mol. Cell. Endocrinol.* **2015**, *416*, 88–96. [CrossRef] [PubMed]

70. Lu, M.H.; Tang, B.; Zeng, S.; Hu, C.J.; Xie, R.; Wu, Y.Y.; Wang, S.M.; He, F.T.; Yang, S.M. Long noncoding RNA BC032469, a novel competing endogenous RNA, upregulates hTERT expression by sponging miR-1207–5p and promotes proliferation in gastric cancer. *Oncogene* **2015**, *413*. [CrossRef]

71. Yang, Z.; Zhi, Q.; Wang, D.; Zhang, L.; Kuang, Y.; Miao, R.; Shi, Y.; Guo, X. Long noncoding RNA C21orF96 promotes the migration, invasion and lymph node metastasis in gastric cancer. *Anticancer Agents. Med. Chem.* **2015**. in press.

72. Vencken, S.F.; Greene, C.M.; McKiernan, P.J. Non-coding RNA as lung disease biomarkers. *Thorax* **2015**, *70*, 501–503. [CrossRef] [PubMed]

73. Van Roosbroeck, K.; Pollet, J.; Calin, G.A. miRNAs and long noncoding RNAs as biomarkers in human diseases. *Expert Rev. Mol. Diagn.* **2013**, *13*, 183–204. [CrossRef] [PubMed]

74. Cortez, M.A.; Bueso-Ramos, C.; Ferdin, J.; Lopez-Berestein, G.; Sood, A.K.; Calin, G.A. MicroRNAs in body fluids—the mix of hormones and biomarkers. *Nat. Rev. Clin. Oncol.* **2011**, *8*, 467–477. [CrossRef] [PubMed]

75. Mall, C.; Rocke, D.M.; Durbin-Johnson, B.; Weiss, R.H. Stability of miRNA in human urine supports its biomarker potential. *Biomark. Med.* **2013**, *7*, 623–631. [CrossRef] [PubMed]

76. Catalucci, D.; Gallo, P.; Condorelli, G. MicroRNAs in Cardiovascular Biology and Heart Disease. *Circ. Cardiovasc. Genet.* **2009**, *2*, 402–408. [CrossRef] [PubMed]

77. Dai, R.; Ahmed, S.A. MicroRNA, a new paradigm for understanding immunoregulation, inflammation, and autoimmune diseases. *Transl. Res.* **2011**, *157*, 163–179. [CrossRef] [PubMed]

78. Clark, M.B.; Johnston, R.L.; Inostroza-Ponta, M.; Fox, A.H.; Fortini, E.; Moscato, P.; Dinger, M.E.; Mattick, J.S. Genome-wide analysis of long noncoding RNA stability. *Genome Res.* **2012**, *22*, 885–898. [CrossRef] [PubMed]

79. Zarate, R.; Boni, V.; Bandres, E.; Garcia-Foncillas, J. MiRNAs and LincRNAs: Could They Be Considered as Biomarkers in Colorectal Cancer? *Int. J. Mol. Sci.* **2012**, *13*, 840–865. [CrossRef] [PubMed]

80. Liao, J.; Yu, L.; Mei, Y.; Guarnera, M.; Shen, J.; Li, R.; Liu, Z.; Jiang, F. Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. *Mol. Cancer* **2010**, *9*, 198. [CrossRef] [PubMed]

81. Zhang, L.; Yang, M.; Marks, P.; White, L.M.; Hurtig, M.; Divine, G.; Gibson, G. Serum non-coding RNAs as biomarkers for osteoarthritis progression after ACL injury. *Osteoarthr. Cartil.* **2012**, *20*, 1631–1637. [CrossRef] [PubMed]

82. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al.* RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science* **2007**, *316*, 1484–1488. [CrossRef] [PubMed]

83. Ponting, C.P.; Oliver, P.L.; Reik, W. Evolution and Functions of Long Noncoding RNAs. *Cell* **2009**, *136*, 629–641. [CrossRef] [PubMed]

84. Wang, K.C.; Yang, Y.W.; Liu, B.; Sanyal, A.; Corces-Zimmerman, R.; Chen, Y.; Lajoie, B.R.; Protacio, A.; Flynn, R.A.; Gupta, R.A.; *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **2011**, *472*, 120–124. [CrossRef] [PubMed]

85. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. *Science* **2008**, *322*, 750–756. [CrossRef] [PubMed]

86. Fitzpatrick, G.V.; Soloway, P.D.; Higgins, M.J. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat. Genet.* **2002**, *32*, 426–431. [CrossRef] [PubMed]

87. Bertani, S.; Sauer, S.; Bolotin, E.; Sauer, F. The non-coding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting Mll1 to chromatin. *Mol. Cell* **2011**, *43*, 1040–1046. [CrossRef] [PubMed]

88. Nagano, T.; Mitchell, J.A.; Sanz, L.A.; Pauler, F.M.; Ferguson-Smith, A.C.; Feil, R.; Fraser, P. The Air Noncoding RNA Epigenetically Silences Transcription by Targeting G9a to Chromatin. *Science* **2008**, *322*, 1717–1720. [CrossRef] [PubMed]

89. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al.* A large intergenic non-coding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419. [CrossRef] [PubMed]

90. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* **2007**, *129*, 1311–1323. [CrossRef] [PubMed]

91. Xing, Z.; Lin, A.; Li, C.; Liang, K.; Wang, S.; Liu, Y.; Park, P.K.; Qin, L.; Wei, Y.; Hawke, D.H.; *et al.* lncRNA Directs Cooperative Epigenetic Regulation Downstream of Chemokine Signals. *Cell* **2014**, *159*, 1110–1125. [CrossRef] [PubMed]

92. Khorkova, O.; Hsiao, J.; Wahlestedt, C. Basic biology and therapeutic implications of lncRNA. *Adv. Drug Deliv. Rev.* **2015**, *87*, 15–24. [CrossRef] [PubMed]

93. Lalevée, S.; Feil, R. Long noncoding RNAs in human disease: Emerging mechanisms and therapeutic strategies. *Epigenomics* **2015**, *7*, 877–879. [CrossRef] [PubMed]

94. Martens, J.A.; Laprade, L.; Winston, F. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature* **2004**, *429*, 571–574. [CrossRef] [PubMed]

95. Hirota, K.; Miyoshi, T.; Kugou, K.; Hoffman, C.S.; Shibata, T.; Ohta, K. Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. *Nature* **2008**, *456*, 130–134. [CrossRef] [PubMed]

96. Martianov, I.; Ramadass, A.; Serra Barros, A.; Chow, N.; Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **2007**, *445*, 666–670. [CrossRef] [PubMed]

97. Wang, X.; Arai, S.; Song, X.; Reichart, D.; Du, K.; Pascual, G.; Tempst, P.; Rosenfeld, M.G.; Glass, C.K.; Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **2008**, *454*, 126–130. [CrossRef] [PubMed]

98. Feng, J.; Bi, C.; Clark, B.S.; Mady, R.; Shah, P.; Kohtz, J.D. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev.* **2006**, *20*, 1470–1484. [CrossRef] [PubMed]

99. Shamovsky, I.; Ivannikov, M.; Kandel, E.S.; Gershon, D.; Nudler, E. RNA-mediated response to heat shock in mammalian cells. *Nature* **2006**, *440*, 556–560. [CrossRef] [PubMed]

100. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A Strategy for Probing the Function of Noncoding RNAs Finds a Repressor of NFAT. *Science* **2005**, *309*, 1570–1573. [CrossRef] [PubMed]

101. Wilusz, J.E.; Freier, S.M.; Spector, D.L. 3′ end processing of a long nuclear-retained non-coding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **2008**, *135*, 919–932. [CrossRef] [PubMed]

102. Zhou, X.; Gao, Q.; Wang, J.; Zhang, X.; Liu, K.; Duan, Z. Linc-RNA-RoR acts as a "sponge" against mediation of the differentiation of endometrial cancer stem cells by microRNA-145. *Gynecol. Oncol.* **2014**, *133*, 333–339. [CrossRef] [PubMed]

103. Modali, S.D.; Parekh, V.I.; Kebebew, E.; Agarwal, S.K. Epigenetic Regulation of the lncRNA MEG3 and Its Target c-MET in Pancreatic Neuroendocrine Tumors. *Mol. Endocrinol.* **2015**, *29*, 224–237. [CrossRef] [PubMed]

104. Butler, A.A.; Webb, W.M.; Lubin, F.D. Regulatory RNAs and control of epigenetic mechanisms: Expectations for cognition and cognitive dysfunction. *Epigenomics* **2015**, *8*, 135–151. [CrossRef] [PubMed]

105. Peschansky, V.J.; Pastori, C.; Zeier, Z.; Motti, D.; Wentzel, K.; Velmeshev, D.; Magistri, M.; Bixby, J.L.; Lemmon, V.P.; Silva, J.P.; *et al.* Changes in expression of the long non-coding RNA FMR4 associate with altered gene expression during differentiation of human neural precursor cells. *Front. Genet.* **2015**, *6*, 263. [CrossRef] [PubMed]

106. Sohi, G.; Dilworth, F.J. Noncoding RNAs as epigenetic mediators of skeletal muscle regeneration. *FEBS J.* **2015**, *282*, 1630–1646. [CrossRef] [PubMed]

107. Pandey, R.R.; Mondal, T.; Mohammad, F.; Enroth, S.; Redrup, L.; Komorowski, J.; Nagano, T.; Mancini-DiNardo, D.; Kanduri, C. Kcnq1ot1 Antisense Noncoding RNA Mediates Lineage-Specific Transcriptional Silencing through Chromatin-Level Regulation. *Mol. Cell* **2008**, *32*, 232–246. [CrossRef] [PubMed]

108. Yoon, J.H.; Abdelmohsen, K.; Gorospe, M. Posttranscriptional Gene Regulation by Long Noncoding RNA. *J. Mol. Biol.* **2013**, *425*, 3723–3730. [CrossRef] [PubMed]

109. Cheetham, S.W.; Gruhl, F.; Mattick, J.S.; Dinger, M.E. Long noncoding RNAs and the genetics of cancer. *Br. J. Cancer* **2013**, *108*, 2419–2425. [CrossRef] [PubMed]

110. Iyer, M.K.; Niknafs, Y.S.; Malik, R.; Singhal, U.; Sahu, A.; Hosono, Y.; Barrette, T.R.; Prensner, J.R.; Evans, J.R.; Zhao, S.; *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **2015**, *47*, 199–208. [CrossRef] [PubMed]

111. Xue, M.; Chen, W.; Li, X. Urothelial cancer associated 1: A long noncoding RNA with a crucial role in cancer. *J. Cancer Res. Clin. Oncol.* **2015**. in press. [CrossRef] [PubMed]

112. Ling, H.; Spizzo, R.; Atlasi, Y.; Nicoloso, M.; Shimizu, M.; Redis, R.S.; Nishida, N.; Gafà, R.; Song, J.; Guo, Z.; *et al.* CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. *Genome Res.* **2013**, *23*, 1446–1461. [CrossRef] [PubMed]

113. Prensner, J.R.; Chen, W.; Han, S.; Iyer, M.K.; Cao, Q.; Kothari, V.; Evans, J.R.; Knudsen, K.E.; Paulsen, M.T.; Ljungman, M.; *et al.* The Long Non-Coding RNA PCAT-1 Promotes Prostate Cancer Cell Proliferation through cMyc. *Neoplasia* **2014**, *16*, 900–908. [CrossRef] [PubMed]

114. Hansji, H.; Leung, E.Y.; Baguley, B.C.; Finlay, G.J.; Askarian-Amiri, M.E. Keeping abreast with long non-coding RNAs in mammary gland development and breast cancer. *Front. Genet.* **2014**, *5*, 379. [CrossRef] [PubMed]

115. Wang, J.; Liu, X.; Wu, H.; Ni, P.; Gu, Z.; Qiao, Y.; Chen, N.; Sun, F.; Fan, Q. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res.* **2010**, *38*, 5366–5383. [CrossRef] [PubMed]

116. Jiang, H.J.; Wang, S.; Ding, Y. Emerging paradigms of long non-coding RNAs in gastrointestinal cancer. *Am. J. Stem Cells* **2014**, *3*, 63–73. [PubMed]

117. Li, H.J.; Li, X.; Pang, H.; Pan, J.J.; Xie, X.J.; Chen, W. Long non-coding RNA UCA1 promotes glutamine metabolism by targeting miR-16 in human bladder cancer. *Jpn. J. Clin. Oncol.* **2015**, *45*, 1055–1063. [CrossRef] [PubMed]

118. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; Laurent, G.S.; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β-secretase expression. *Nat. Med.* **2008**, *14*, 723–730. [CrossRef] [PubMed]

119. Mus, E.; Hof, P.R.; Tiedge, H. Dendritic BC200 RNA in aging and in Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10679–10684. [CrossRef] [PubMed]

120. Sandhya, P.; Joshi, K.; Scaria, V. Long noncoding RNAs could be potential key players in the pathophysiology of Sjögren's syndrome. *Int. J. Rheum. Dis.* **2015**, *18*, 898–905. [CrossRef] [PubMed]

121. Carpenter, S. Long noncoding RNA: Novel links between gene expression and innate immunity. *Virus Res.* **2015**, *15*, 30051–30054. [CrossRef] [PubMed]

122. Pan, Q.; Shai, O.; Lee, L.J.; Frey, B.J.; Blencowe, B.J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **2008**, *40*, 1413–1415. [CrossRef] [PubMed]

123. Kapranov, P.; Drenkow, J.; Cheng, J.; Long, J.; Helt, G.; Dike, S.; Gingeras, T.R. Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res.* **2005**, *15*, 987–997. [CrossRef] [PubMed]

124. Salzman, J.; Gawad, C.; Wang, P.L.; Lacayo, N.; Brown, P.O. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS ONE* **2012**, *7*, e30733. [CrossRef] [PubMed]

125. Salgia, S.R.; Singh, S.K.; Gurha, P.; Gupta, R. Two reactions of Haloferax volcanii RNA splicing enzymes: Joining of exons and circularization of introns. *RNA* **2003**, *9*, 319–330. [CrossRef] [PubMed]

126. Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Kinzler, K.W.; Vogelstein, B. Scrambled exons. *Cell* **1991**, *64*, 607–613. [CrossRef]

127. Yin, Q.F.; Yang, L.; Zhang, Y.; Xiang, J.F.; Wu, Y.W.; Carmichael, G.G.; Chen, L.L. Long Noncoding RNAs with snoRNA Ends. *Mol. Cell* **2012**, *48*, 219–230. [CrossRef] [PubMed]

128. Zhang, Y.; Zhang, X.O.; Chen, T.; Xiang, J.F.; Yin, Q.F.; Xing, Y.H.; Zhu, S.; Yang, L.; Chen, L.L. Circular Intronic Long Noncoding RNAs. *Mol. Cell* **2013**, *51*, 792–806. [CrossRef] [PubMed]

129. Wilusz, J.E.; JnBaptiste, C.K.; Lu, L.Y.; Kuhn, C.D.; Joshua-Tor, L.; Sharp, P.A. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **2012**, *26*, 2392–2407. [CrossRef] [PubMed]

130. Wilusz, J.E.; Whipple, J.M.; Phizicky, E.M.; Sharp, P.A. tRNAs marked with CCACCA are targeted for degradation. *Science* **2011**, *334*, 817–821. [CrossRef] [PubMed]

131. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388. [CrossRef] [PubMed]

132. Zhang, X.; Rice, K.; Wang, Y.; Chen, W.; Zhong, Y.; Nakayama, Y.; Zhou, Y.; Klibanski, A. Maternally Expressed Gene 3 (MEG3) Noncoding Ribonucleic Acid: Isoform Structure, Expression, and Functions. *Endocrinology* **2010**, *151*, 939–947. [CrossRef] [PubMed]

133. Johnsson, P.; Ackley, A.; Vidarsdottir, L.; Lui, W.O.; Corcoran, M.; Grandér, D.; Morris, K.V. A pseudogene long noncoding RNA network regulates PTEN transcription and translation in human cells. *Nat. Struct. Mol. Biol.* **2013**, *20*, 440–446. [CrossRef] [PubMed]

134. Borsani, G.; Tonlorenzi, R.; Simmler, M.C.; Dandolo, L.; Arnaud, D.; Capra, V.; Grompe, M.; Pizzuti, A.; Muzny, D.; Lawrence, C.; *et al.* Characterization of a murine gene expressed from the inactive X chromosome. *Nature* **1991**, *351*, 325–329. [CrossRef] [PubMed]

135. Brockdorff, N.; Ashworth, A.; Kay, G.F.; Cooper, P.; Smith, S.; McCabe, V.M.; Norris, D.P.; Penny, G.D.; Patel, D.; Rastan, S. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **1991**, *351*, 329–331. [CrossRef] [PubMed]

136. Brown, C.J.; Ballabio, A.; Rupert, J.L.; Lafreniere, R.G.; Grompe, M.; Tonlorenzi, R.; Willard, H.F. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **1991**, *349*, 38–44. [CrossRef] [PubMed]

137. Burd, C.E.; Jeck, W.R.; Liu, Y.; Sanoff, H.K.; Wang, Z.; Sharpless, N.E. Expression of Linear and Novel Circular Forms of an INK4/ARF-Associated Non-Coding RNA Correlates with Atherosclerosis Risk. *PLoS Genet.* **2010**, *6*, e1001233. [CrossRef] [PubMed]

138. Ji, P.; Diederichs, S.; Wang, W.B.; Boing, S.; Metzger, R.; Schneider, P.M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; *et al.* MALAT-1, a novel noncoding RNA, and thymosin beta 4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **2003**, *22*, 8031–8041. [CrossRef] [PubMed]

139. Xu, C.; Yang, M.; Tian, J.; Wang, X.; Li, Z. MALAT-1: A long non-coding RNA and its important 3′ end functional motif in colorectal cancer metastasis. *Int. J. Oncol.* **2011**, *39*, 169–175. [PubMed]

140. Davis, I.J.; Hsi, B.L.; Arroyo, J.D.; Vargas, S.O.; Yeh, Y.A.; Motyckova, G.; Valencia, P.; Perez-Atayde, A.R.; Argani, P.; Ladanyi, M.; *et al*. Cloning of an Alpha-TFEB fusion in renal tumors harboring the t(6;11)(p21;q13) chromosome translocation. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6051–6056. [CrossRef] [PubMed]

141. Brown, J.A.; Valenstein, M.L.; Yario, T.A.; Tycowski, K.T.; Steitz, J.A. Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19202–19207. [CrossRef] [PubMed]

142. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA Gas5 Is a Growth Arrest and Starvation-Associated Repressor of the Glucocorticoid Receptor. *Sci. Signal.* **2010**, *3*, ra8. [CrossRef] [PubMed]

143. Novikova, I.V.; Hennelly, S.P.; Sanbonmatsu, K.Y. Sizing up long non-coding RNAs Do lncRNAs have secondary and tertiary structure? *BioArchitecture* **2012**, *2*, 189–199. [CrossRef] [PubMed]

144. Weeks, K.M. Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.* **2010**, *20*, 295–304. [CrossRef] [PubMed]

145. Harris, M.E.; Christian, E.L. RNA Crosslinking Methods. *Methods Enzymol.* **2009**, *468*, 127–146. [PubMed]

146. Tullius, T.D.; Greenbaum, J.A. Mapping nucleic acid structure by hydroxyl radical cleavage. *Curr. Opin. Chem. Biol.* **2005**, *9*, 127–134. [CrossRef] [PubMed]

147. Ban, N.; Nissen, P.; Hansen, J.; Moore, P.B.; Steitz, T.A. The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. *Science* **2000**, *289*, 905–920. [CrossRef] [PubMed]

148. Ben-Shem, A.; Garreau de Loubresse, N.; Melnikov, S.; Jenner, L.; Yusupova, G.; Yusupov, M. The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. *Science* **2011**, *334*, 1524–1529. [CrossRef] [PubMed]

149. Brion, P.; Westhof, E. Hierachy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113–137. [CrossRef] [PubMed]

150. Sprinzl, M.; Vassilenko, K.S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **2005**, *33*, D139–D140. [CrossRef] [PubMed]

151. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964. [CrossRef] [PubMed]

152. Eddy, S.R.; Durbin, R. RNA sequence analysis using covariance models. *Nucleic Acids Res.* **1994**, *22*, 2079–2088. [CrossRef] [PubMed]

153. Gorodkin, J.; Heyer, L.J.; Stormo, G.D. Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res.* **1997**, *25*, 3724–3732. [CrossRef] [PubMed]

154. Gorodkin, J.; Stricklin, S.L.; Stormo, G.D. Discovering common stemloop motifs in unaligned RNA sequences. *Nucleic Acids Res.* **2001**, *29*, 2135–2144. [CrossRef] [PubMed]

155. Matthews, B.W. comparison of the predicted and observed structure of T4 phage lysozyme. *Biochim. Biophys Acta* **1975**, *405*, 442–451. [CrossRef]

156. Foldalign: RNA Structure and Sequence Alignment. Available online: http://rth.dk/resources/foldalign/ (accessed on 8 October 2015).

157. Mathews, D.H.; Turner, D.H. Dynalign: An algorithm for finding the secondary structure common to two RNA sequences1. *J. Mol. Biol.* **2002**, *317*, 191–203. [CrossRef] [PubMed]

158. Knapp, G. Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol.* **1989**, *180*, 192–212. [PubMed]

159. Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J.P.; Ehresmann, B. Probing the structure of RNAs in solution. *Nucleic Acids Res.* **1987**, *15*, 9109–9128. [CrossRef] [PubMed]

160. Mathews, D.H. Predicting a set of minimal free energy RNA secondary structures common to two sequences. *Bioinformatics* **2005**, *21*, 2246–2253. [CrossRef] [PubMed]

161. Grate, L. Automatic RNA secondary structure determination with stochastic context-free grammars. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1995**, *3*, 136–144. [PubMed]

162. Knudsen, B.; Hein, J. RNA secondary structure prediction using stochastic context-free grammars and evolutionary history. *Bioinformatics* **1999**, *15*, 446–454. [CrossRef] [PubMed]

163. Pfold. Available online: http://www.daimi.au.dk/~compbio/pfold (accessed on 8 October 2015).

164. Hofacker, I.L.; Fekete, M.; Stadler, P.F. Secondary Structure Prediction for Aligned RNA Sequences. *J. Mol. Biol.* **2002**, *319*, 1059–1066. [CrossRef]

165. Hofacker, I.L. Vienna RNA secondary structure server. *Nucleic Acids Res.* **2003**, *31*, 3429–3431. [CrossRef] [PubMed]

166. Waterman, M.S.; Smith, T.F. RNA secondary structure: A complete mathematical analysis. *Math. Biosci.* **1978**, *42*, 257–266. [CrossRef]

167. The ViennaRNA Web Services. Available online: http://rna.tbi.univie.ac.at (accessed on 8 October 2015).

168. Siebert, S.; Backofen, R. MARNA: Multiple alignment and consensus structure prediction of RNAs based on sequence structure comparisons. *Bioinformatics* **2005**, *21*, 3352–3359. [CrossRef] [PubMed]

169. Corpet, F.; Michot, B. RNAlign program: Alignment of RNA sequences using both primary and secondary structures. *Comput. Appl. Biosci.* **1994**, *10*, 389–399. [CrossRef] [PubMed]

170. MARNA-Structure Alignment. Available online: http://rna.informatik.uni-freiburg.de/MARNA/Input.jsp (accessed on 8 October 2015).

171. Norman, R.P.; Brian, C.T.; Carl, R.W. *The RNA World*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, USA, 1993; pp. 91–117.

172. Mathews, D.H. Revolutions in RNA secondary structure prediction. *J. Mol. Biol.* **2006**, *359*, 526–532. [CrossRef] [PubMed]

173. Zuker, M. On finding all suboptimal foldings of an RNA molecule. *Science* **1989**, *244*, 48–52. [CrossRef] [PubMed]

174. Buratti, E.; Baralle, F.E. Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell. Biol.* **2004**, *24*, 10505–10514. [CrossRef] [PubMed]

175. Yang, Y.; Zhan, L.; Zhang, W.; Sun, F.; Wang, W.; Tian, N.; Bi, J.; Wang, H.; Shi, D.; Jiang, Y.; *et al.* RNA secondary structure in mutually exclusive splicing. *Nat. Struct. Mol. Biol.* **2011**, *18*, 159–168. [CrossRef] [PubMed]

176. The mfold Web Server. Available online: http://unafold.rna.albany.edu/?q=mfold (accessed on 8 October 2015).

177. Humann, F.C.; Tiberio, G.J.; Hartfelder, K. Sequence and expression characteristics of long noncoding RNAs in honey bee caste development–potential novel regulators for transgressive ovary size. *PLoS ONE* **2013**, *8*, e78915. [CrossRef] [PubMed]

178. RNAfold web server. Available online: http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi. (accessed on 8 October 2015).

179. Giegerich, R.; Voss, B.; Rehmsmeier, M. Abstract shapes of RNA. *Nucleic Acids Res.* **2004**, *32*, 4843–4851. [CrossRef] [PubMed]

180. RNAshapes. Available online: http://bibiserv.techfak.uni-bielefeld.de/rnashapes (accessed on 8 October 2015).

181. Xia, T.; SantaLucia, J.; Burkard, M.E.; Kierzek, R.; Schroeder, S.J.; Jiao, X.; Cox, C.; Turner, D.H. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson−Crick Base Pairs. *Biochemistry* **1998**, *37*, 14719–14735. [CrossRef] [PubMed]

182. Mathews, D.H.; Disney, M.D.; Childs, J.L.; Schroeder, S.J.; Zuker, M.; Turner, D.H. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7287–7292. [CrossRef] [PubMed]

183. Ding, Y.; Tang, Y.; Kwok, C.K.; Zhang, Y.; Bevilacqua, P.C.; Assmann, S.M. *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **2014**, *505*, 696–700. [CrossRef] [PubMed]

184. Khaitan, D.; Dinger, M.E.; Mazar, J.; Crawford, J.; Smith, M.A.; Mattick, J.S.; Perera, R.J. The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. *Cancer Res.* **2011**, *71*, 3852–3862. [CrossRef] [PubMed]

185. RNAstructure, Version 5.8. Available online: http://rna.urmc.rochester.edu/RNAstructure.html (accessed on 8 October 2015).

186. Shapiro, B.A.; Wu, J.C.; Bengali, D.; Potts, M.J. The massively parallel genetic algorithm for RNA folding: MIMD implementation and population variation. *Bioinformatics* **2001**, *17*, 137–148. [CrossRef] [PubMed]

187. Ding, Y.; Lawrence, C.E. Statistical prediction of single-stranded regions in RNA secondary structure and application to predicting effective antisense target sites and beyond. *Nucleic Acids Res.* **2001**, *29*, 1034–1046. [CrossRef] [PubMed]

188. Do, C.B.; Woods, D.A.; Batzoglou, S. CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics* **2006**, *22*, E90–E98. [CrossRef] [PubMed]

189. Dann Iii, C.E.; Wakeman, C.A.; Sieling, C.L.; Baker, S.C.; Irnov, I.; Winkler, W.C. Structure and Mechanism of a Metal-Sensing Regulatory RNA. *Cell* **2007**, *130*, 878–892. [CrossRef] [PubMed]

190. Das, R.; Baker, D. Automated de novo prediction of native-like RNA tertiary structures. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14664–14669. [CrossRef] [PubMed]

191. Yarov-Yarovoy, V.; Schonbrun, J.; Baker, D. Multipass Membrane Protein Structure Prediction Using Rosetta. *Proteins* **2006**, *62*, 1010–1025. [CrossRef] [PubMed]

192. Hajdin, C.E.; Ding, F.; Dokholyan, N.V.; Weeks, K.M. On the significance of an RNA tertiary structure prediction. *RNA* **2010**, *16*, 1340–1349. [CrossRef] [PubMed]

193. Das, R.; Kudaravalli, M.; Jonikas, M.; Laederach, A.; Fong, R.; Schwans, J.P.; Baker, D.; Piccirilli, J.A.; Altman, R.B.; Herschlag, D. Structural inference of native and partially folded RNA by high-throughput contact mapping. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4144–4149. [CrossRef] [PubMed]

194. Das, R.; Karanicolas, J.; Baker, D. Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* **2010**, *7*, 291–294. [CrossRef] [PubMed]

195. Jonikas, M.A.; Radmer, R.J.; Laederach, A.; Das, R.; Pearlman, S.; Herschlag, D.; Altman, R.B. Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA* **2009**, *15*, 189–199. [CrossRef] [PubMed]

196. NAST (The Nucleic Acid Simulation Tool). Available online: https://simtk.org/home/nast (accessed on 8 October 2015).

197. Ding, F.; Sharma, S.; Chalasani, P.; Demidov, V.V.; Broude, N.E.; Dokholyan, N.V. Ab initio RNA folding by discrete molecular dynamics: From structure prediction to folding mechanisms. *RNA* **2008**, *14*, 1164–1173. [CrossRef] [PubMed]

198. Ding, F.; Dokholyan, N.V. Emergence of Protein Fold Families through Rational Design. *PLoS Comput. Biol.* **2006**, *2*, e85. [CrossRef] [PubMed]

199. Gherghe, C.M.; Leonard, C.W.; Ding, F.; Dokholyan, N.V.; Weeks, K.M. Native-like RNA Tertiary Structures Using a Sequence-Encoded Cleavage Agent and Refinement by Discrete Molecular Dynamics. *J. Am. Chem. Soc.* **2009**, *131*, 2541–2546. [CrossRef] [PubMed]

200. Frellsen, J.; Moltke, I.; Thiim, M.; Mardia, K.V.; Ferkinghoff-Borg, J.; Hamelryck, T. A probabilistic model of RNA conformational space. *PLoS Comput. Biol.* **2009**, *5*, e1000406. [CrossRef] [PubMed]

201. Xia, Z.; Gardner, D.P.; Gutell, R.R.; Ren, P. Coarse-Grained Model for Simulation of RNA Three-Dimensional Structures. *J. Phys. Chem. B* **2010**, *114*, 13497–13506. [CrossRef] [PubMed]

202. Gutell, R.R.; Lee, J.C.; Cannone, J.J. The accuracy of ribosomal RNA comparative structure models. *Curr. Opin. Struct. Biol.* **2002**, *12*, 301–310. [CrossRef]

203. Martinez, H.M.; Maizel, J.J.; Shapiro, B.A. RNA2D3D: A program for generating, viewing, and comparing 3-dimensional models of RNA. *J. Biomol. Struct. Dyn.* **2008**, *25*, 669–683. [CrossRef] [PubMed]

204. Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham Iii, T.E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41. [CrossRef]

205. Yingling, Y.G.; Shapiro, B.A. Dynamic behavior of the telomerase RNA hairpin structure and its relationship to dyskeratosis congenita. *J. Mol. Biol.* **2005**, *348*, 27–42. [CrossRef] [PubMed]

206. Yingling, Y.G.; Shapiro, B.A. The impact of dyskeratosis congenita mutations on the structure and dynamics of the human telomerase RNA pseudoknot domain. *J. Biomol. Struct. Dyn.* **2007**, *24*, 303–320. [CrossRef] [PubMed]

207. Yingling, Y.G.; Shapiro, B.A. The prediction of the wild-type telomerase RNA pseudoknot structure and the pivotal role of the bulge in its formation. *J. Mol. Graph. Model.* **2006**, *25*, 261–274. [CrossRef] [PubMed]

208. Cao, S.; Chen, S.J. Physics-based de novo prediction of RNA 3D structures. *J. Phys. Chem. B* **2011**, *115*, 4216–4226. [CrossRef] [PubMed]

209. RSIMS. Available online: https://github.com/jpbida/rsim (accessed on 8 October 2015).

210. Zhao, Y.; Gong, Z.; Xiao, Y. Improvements of the hierarchical approach for predicting RNA tertiary structure. *J. Biomol. Struct. Dyn.* **2011**, *28*, 815–826. [CrossRef] [PubMed]

211. 3dRNA: Automatic building of three-dimensional RNA structures. Available online: http://biophy. hust.edu.cn/3dRNA/3dRNA-1.0.html (accessed on 8 October 2015).

212. Parisien, M.; Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **2008**, *452*, 51–55. [CrossRef] [PubMed]

213. Jossinet, F.; Ludwig, T.E.; Westhof, E. Assemble: An interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinformatics* **2010**, *26*, 2057–2059. [CrossRef] [PubMed]

214. Zhao, Y.; Wang, J.; Chen, X.; Luo, H.; Zhao, Y.; Xiao, Y.; Chen, R. Large-scale study of long non-coding RNA functions based on structure and expression features. *Sci. China Life Sci.* **2013**, *56*, 953–959. [CrossRef] [PubMed]

215. Laing, C.; Schlick, T. Computational approaches to RNA structure prediction, analysis, and design. *Curr. Opin. Struct. Biol.* **2011**, *21*, 306–318. [CrossRef] [PubMed]

216. Laing, C.; Schlick, T. Computational approaches to 3D modeling of RNA. *J. Phys. Condens. Matter* **2010**, *283101*. [CrossRef] [PubMed]

217. Wapinski, O.; Chang, H.Y. Long noncoding RNAs and human disease. *Trends Cell Biol.* **2011**, *21*, 354–361. [CrossRef] [PubMed]

218. Lyngso, R.B.; Pedersen, C.N. RNA pseudoknot prediction in energy-based models. *J. Comput. Biol.* **2000**, *7*, 409–427. [CrossRef] [PubMed]

219. Mathews, D.H.; Sabina, J.; Zuker, M.; Turner, D.H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure1. *J. Mol. Biol.* **1999**, *288*, 911–940. [CrossRef] [PubMed]